



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/631,172	07/31/2003	James Michael McArdle	AUS920030382US1	1419
45993	7590	05/12/2008		
IBM CORPORATION (RHF) C/O ROBERT H. FRANTZ P. O. BOX 23324 OKLAHOMA CITY, OK 73123			EXAMINER HOFFLER, RAHEEM	
			ART UNIT 2165	PAPER NUMBER
			MAIL DATE 05/12/2008	DELIVERY MODE PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No. 10/631,172	Applicant(s) MCARDLE, JAMES MICHAEL	
	Examiner RAHEEM HOFFLER	Art Unit 2165	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 13 February 2008.
- 2a) ☒ This action is **FINAL**. 2b) ☐ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-5 and 10-18 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 1-5 and 10-18 is/are rejected.
- 7) ☐ Claim(s) _____ is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☒ The drawing(s) filed on 31 July 2003 is/are: a) ☒ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
 2. ☐ Certified copies of the priority documents have been received in Application No. _____.
 3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- | | |
|--|---|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892) | 4) <input type="checkbox"/> Interview Summary (PTO-413) |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948) | Paper No(s)/Mail Date. _____ |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO/SB/08) | 5) <input type="checkbox"/> Notice of Informal Patent Application |
| Paper No(s)/Mail Date _____ | 6) <input type="checkbox"/> Other: _____ |

Detailed Action

Response to Amendment

The Office Action has been issued in response to amendment filed 31 August 2007. Claims 1-5 & 10-18 are pending. Applicant's arguments have been carefully and respectfully considered in light of the instant amendment, and are persuasive. Accordingly, this action has been made Non-FINAL.

Claim Rejections - 35 USC § 112

The following is a quotation of the second paragraph of 35 U.S.C. 112:

The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

Claims 1, 10 and 15 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

Claim 1 makes reference to "the modified fields" but fails to disclose what "the modified fields" are, thus, making the referenced term indefinite.

Claim Rejections - 35 USC § 101

35 U.S.C. 101 reads as follows:

Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

Claims 15-18 of the claimed invention is directed to non-statutory subject matter. Claim 15 is a system claim comprising a number of functions, but fail to include any type of hardware such as a processor in which these functions can perform. Thus, claim 15 is considered software, per se.

Art Unit: 2165

Software, *per se*:

The claims lack the necessary physical articles or objects to constitute a machine or a manufacture within the meaning of 35 USC 101. They are clearly not a series of steps or acts to be a process nor are they a combination of chemical compounds to be a composition of matter. As such, they fail to fall within a statutory category. They are, at best, functional descriptive material *per se*.

Descriptive material can be characterized as either “functional descriptive material” or “nonfunctional descriptive material.” Both types of “descriptive material” are nonstatutory when claimed as descriptive material *per se*, 33 F.3d at 1360, 31 USPQ2d at 1759. When functional descriptive material is recorded on some computer-readable medium, it becomes structurally and functionally interrelated to the medium and will be statutory in most cases since use of technology permits the function of the descriptive material to be realized. Compare *In re Lowry*, 32 F.3d 1579, 1583-84, 32 USPQ2d 1031, 1035 (Fed. Cir. 1994)

Merely claiming nonfunctional descriptive material, i.e., abstract ideas, stored on a computer-readable medium, in a computer, or on an electromagnetic carrier signal, does not make it statutory. See *Diehr*, 450 U.S. at 185-86, 209 USPQ at 8 (noting that the claims for an algorithm in *Benson* were unpatentable as abstract ideas because “[t]he sole practical application of the algorithm was in connection with the programming of a general purpose computer.”).

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

Claims 1-5 & 10-18 are rejected under 35 U.S.C. 103(a) as being unpatentable over Burdick et al (USPG Pub No. 20040107203A1; Burdick hereinafter) in view of Applicant admitted prior art (Background: Figure 1 & 2; see paragraphs [0001-0031]; Background hereinafter)

As for Claim 1, Burdick et al teaches, “declaring said data feature as suspect responsive to said degree of correlation exceeding a threshold” (see paragraph [0035], [0053]), “generating a set of cleaning attributes for each cleaned data record in a

complete set of cleaned data records, said cleaning attributes reflecting which fields of each record have been modified by a cleaning operation” (see Fig. 1; see paragraph [0038]; e.g., attribute of an entity), and “determining a degree of correlation of said data feature to the modified fields of said subset of cleaned data records according to said cleaning attributes” (see paragraph [0035]; e.g., determining if two records are duplicates involves performing a similarity test that qualifies the similarity of two records), but fails to explicitly recite, “...receiving a data feature identified by a data mining process for a subset...”.

Background explicitly recites, “receiving a data feature identified by a data mining process for a subset of said complete set of cleaned data records” (see paragraph [0010-0017; e.g., utilizing data mining tools and techniques to return particular information from said complete set of cleaned data).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to have combined the method and system of data analysis taught within Applicant’s background with the data cleansing application of Burdick in order to produce much more accurate and efficient results than could be obtained previously (see Burdick [0008]).

As for Claim 2, Burdick teaches, forming “...a cleaning attributes register for each cleaned data record” (see paragraph [0057-0058]), and “generating a set of bit-mapped Boolean flags” (see paragraph [0058]; e.g., Boolean expressions utilized for the data

cleaning process.)

As for Claim 3, Burdick teaches, “performing an operation selected from the group of appending a set of cleaning attributes to each cleaned data record, prepending a set of cleaning attributes to each cleaned data record, distributing a set of cleaning attributes to each cleaned data record, and generating a cleaning attribute table” (see Figure 1; see paragraph [0034], [0038]; e.g., attributes of an entity).

As for Claim 4, Burdick teaches, “a step selected from the group of receiving a cluster, receiving a trend, and receiving a pattern” (see paragraph [0026], [0032]; e.g., clustering, matching, standardization).

As for Claim 5, Burdick teaches, “comparing each record in a raw data set to each record in a cleaned data set” (see paragraph [0069-0070]).

Claims 10-14 differ from Claims 1-5 in that claims 10-14 are computer readable medium whereas claims 1-5 are method claims. Thus, claims 10-14 are analyzed as previously discussed with respect to claims 1-5 above.

Claims 15-18 differ from Claims 1-4 in that claims 15-18 are system whereas claims 1-4 are method claims. Thus, claims 15-18 are analyzed as previously

discussed with respect to claims 1-4 above.

Response to Arguments

Applicant's arguments with respect to claims 1-5 & 10-18 have been carefully considered but are not persuasive in view of the original grounds of rejection.

With respect to Applicant's argument that:

"Claim1. We agree with the Examiner's finding that Burdick fails to teach our claim elements of:

"generating a set of cleaning attributes for each cleaned data record in a complete set of cleaned data records, said cleaning attributes reflecting which fields of each record have been modified by a cleaning operation";

"receiving a data feature identified by a data mining process for a subset of said complete set of cleaned data records";

"determining a degree of correlation of said data feature to the modified fields of said subset of cleaned data records according to said cleaning attributes"...

We believe that one of ordinary skill in the art applying plain meanings of the terms used in Wocke's disclosure would read their after Grow attribute to be a single attribute associated with an entire set of data records and all of the fields within the data records.

We do not believe that it would have been obvious to one of ordinary skill in the art to further modify Wocke, or Burdick, to record a cleansing *attribute for each field in each record* of a set of data records. If one uses Wocke and Burdick as indicators of the ordinary skill level in the art at the time of filing our application, then it is noticeable that neither Burdick or Wocke recognize the problem of relying upon data features in subsets of data records which include cleansed data fields. Wocke recognizes a need to generally know if a data set has been grown, but does not recognize the need to further identify *which fields* within that data set were grown (or otherwise modified) so that any data features including those modified fields could be considered *suspect* (e.g. given less or no weight in decision making)."

Examiner is not persuaded. The above argument is not persuasive because the relied upon references of Burdick in combination with Wocke do teach of ...

With respect to Applicant's argument that:

Art Unit: 2165

"Claim 2. We respectfully disagree that Wocke's ¶0045 "symbol meaning" and ¶0121-0124 disclose "bit-mapped" cleansing attributes for each field in each record. ¶0045 contains a list of multi-bit values, thus it is not bit-mapped. And, ¶0124 discloses their single afterGrow flag for the entire "map" (e.g. for the entire output set of data records), and other multi-bit values (e.g. maximum number of iterations, minimum quantization threshold, current step number, etc.), which is not a bit-map."

Examiner is not persuaded. The above argument is moot in view of new grounds of rejection.

With respect to Applicant's argument that:

"Claim 3. We disagree that Wocke discloses appending or prepending our bit-mapped set of cleaning attributes to each record within the data set. We believe Wocke only discloses one afterGrow flag for the entire set of records, thus it is not appended to each record, but instead is associated with the entire set of records."

Examiner is not persuaded. The above argument is moot in view of new grounds of rejection.

With respect to Applicant's argument that:

"Claim 4. We agree that Wocke discloses receiving or finding a cluster as a data feature. However, Claim 4 depends from Claim 1, which recites steps, elements and features (e.g. one cleansing attribute per field in each record), which are not taught or suggested by Burdick or Wocke."

Examiner is not persuaded. The above argument is not persuasive because the relied upon reference of Burdick does teach of the data feature of "clustering" at paragraphs [0026] & [0032] along with teaching additional limitations mentioned in Claim 1 such as "one cleansing attribute per field in each record", as illustrated within Figure 1. Examiner maintains rejection.

With respect to Applicant's argument that:

Claim 5. Claim 5 depends from Claim 1, which recites steps, elements and features (e.g. one cleansing attribute *per field in each record*), which are not taught or suggested by Burdick or Wocke."

Examiner is not persuaded. The above argument is moot in view of new grounds of rejection.

Conclusion

The prior art made of reference and not relied upon is considered pertinent to Applicant's disclosure.

Hetherington et al (USPG Pub No. 20020010714A1) teaches a method and apparatus for processing free-format data.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to RAHEEM HOFFLER whose telephone number is (571)270-1036. The examiner can normally be reached on 7:30 a.m. - 5:00 p.m..

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Christian Chace can be reached on (571) 272-4190. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Art Unit: 2165

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Raheem Hoffer/

Examiner, Art Unit 2165

/H. Q. P./

Primary Examiner, Art Unit 2168

/Christian P. Chace/

Supervisory Patent Examiner, Art Unit 2165